

- (19) Japanese Patent Office
(11) Laid-open Patent Application
(12) Japanese Laid-open Patent Application Publication
No. Sho63-25700 (JP-A-63-25700)
(51)
Int. Cl.⁴ Classification Symbol JPO file number
G10L 9/14 M-8622-5D
(43) Published: February 3, 1988
Request for Examination: No
Number of Claims: 1
(4 pages in total)
(54) Title of the Invention: CONCATENATION METHOD OF PROLONGED
SOUND
(21) Patent Application No. Sho61-168134
(22) Applied: July 18, 1986
(71) Applicant: Matsushita Electric Industrial Co., Ltd.
1006 Oaza-kadoma, Kadoma-shi, Osaka
(72) Inventor: Rieko KOBAYASHI
c/o Matsushita Communication Industrial Co., Ltd.
3-1, Tsunashima-higashi 4-chome, Kohoku-ku, Yokohama-shi,
Kanagawa-ken
(74) Agent: Tsuneji HOSHINO (Patent Attorney) et al.

THIS PAGE BLANK (USPTO)

Specification

1. Title of the Invention: CONCATENATION METHOD OF PROLONGED SOUND

2. What is claimed is:

A concatenation method of a prolonged sound where a subsequent syllable is a prolonged sound, comprising:

in a case where amp parameters are used, concatenating parameters of a preceding syllable until a terminating frame of a prolonged sound section syllable at a constant value; and

in a case where linear predictive coefficient parameters are used, using for a prolonged sound section parameters within a syllable file for an independently uttered vowel same as a vowel of a preceding syllable, and concatenating by linear interpolation these parameters to parameters of a terminating frame of a static segment within a syllable file for the vowel of the preceding syllable.

3. Detailed Description of the Invention

[Industrial Field of the Invention]

This invention relates to a concatenation method of a prolonged sound, in speech synthesis methods using rule-based synthesis.

[Prior Art]

A conventional concatenation method of a prolonged sound is described with reference to Fig. 4. Fig. 4 shows a conventional concatenation method of a prolonged sound using amp parameters. In the figure, reference numeral 31 denotes the duration of a preceding syllable, 32 the duration of a prolonged sound, 33 shows the duration of a subsequent syllable, 34 a vowel segment of the preceding syllable, 35 a parameter for the starting frame of a static segment in the syllable file for the preceding syllable, and 36 a parameter for the terminating frame of the static segment in the syllable file for the preceding syllable. Reference numeral 37 denotes a frame into which an independently uttered vowel same as the vowel in the preceding syllable is embedded, 38 (shaded area) parameters from the head of the syllable file

THIS PAGE BLANK (USPTO)

of the independently uttered vowel to the terminating frame 39 of the static segment, and 39 and 40 the starting and terminating frames of the static segment in the syllable file of the independently uttered vowel. That is to say, the conventional concatenation method of a prolonged sound uses the parameter from the syllable file of the independently uttered vowel same as the vowel in the preceding syllable for the section corresponding to the prolonged sound, and links these parameters with parameters in the syllable file for the preceding syllable, as is the case with concatenation methods of other syllables.

In this way, in the conventional concatenation method of a prolonged sound, when an independently uttered vowel of a preceding syllable is substituted into a section corresponding to a prolonged sound symbol, it is possible to perform processing using an algorithm similar to those for other ordinary syllables. Furthermore, while Fig. 4 shows a case of amp parameters, linear predictive coefficient parameters may also be used in a similar way.

[Problem to be Solved by the Invention]

However, in the conventional method above, an independently uttered vowel same as the vowel in the preceding syllable is used in the section corresponding to a prolonged sound, resulting in the problem of the section not perceptually sounding like a prolonged sound.

This invention provides a concatenation method of a prolonged sound that can synthesize prolonged sounds using rule-based synthesis.

[Means for Solving the Problem]

In order to solve the above problem, this invention provides a concatenation method of a prolonged sound where a subsequent syllable is a prolonged sound, comprising: in a case where amp parameters are used, concatenating parameters of a preceding syllable until a terminating frame of a prolonged sound section syllable at a constant value; and, in a case where linear predictive coefficient parameters are used, using for a prolonged sound section parameters within a syllable file for an

THIS PAGE BLANK (USPTO)

independently uttered vowel same as a vowel of a preceding syllable, and concatenating by linear interpolation these parameters with parameters of a terminating frame of a static segment within a syllable file for the vowel of the preceding syllable..

[Function]

The above method of this invention allows creating a parameter time series required for synthesizing synthesized sounds, containing prolonged sounds, and obtaining a prolonged sound that sounds perceptually natural, through the use of parameters in existing syllable files without adding new prolonged sound files to the syllable files.

[Embodiment]

An embodiment of this invention is described with reference to Fig. 1 and Fig. 2. Fig. 1 shows a concatenation method using amp parameters in an embodiment of the invention, while Fig. 2 shows a concatenation method using linear predictive coefficient parameters in an embodiment of the invention. In Fig. 1, reference numeral 1 denotes the duration of a preceding syllable, 2 the duration of a prolonged sound, 3 the duration of a subsequent syllable, 4 a vowel segment of the preceding syllable, 5 a parameter for the starting frame of a static segment in the syllable file for the preceding syllable, 6 a parameter for the terminating frame of the static segment in the syllable file for the preceding syllable, 7 a section corresponding to the prolonged sound, 8 (shown by a point on the time axis) a parameter for the starting frame of a static segment in the syllable file for the prolonged sound section, and 9 a parameter for the terminating frame of the static segment in the syllable file for the prolonged sound section. As shown in Fig. 1, in the method using amp parameters, the parameter for the terminating frame of the preceding syllable 6 is used as a constant value until the terminating frame of the static segment in the syllable file for the prolonged sound section 9.

In Fig. 2, which shows a concatenation method using linear predictive coefficient parameters, reference numeral 11 denotes

THIS PAGE BLANK (USPTO)

the duration of a preceding syllable, 12 the duration of a prolonged sound, 13 the duration of a subsequent syllable, 14 a vowel segment of the preceding syllable, 15 a parameter for the starting frame of a static segment in the syllable file for the preceding syllable, 16 a parameter for the terminating frame of the static segment in the syllable file for the preceding syllable, 17 a prolonged sound section, 18 a parameter for the starting frame of a static segment in the syllable file for an independently uttered vowel same as the vowel in the preceding syllable, and 19 a parameter for the terminating frame of the static segment in the syllable file for the independently uttered vowel.

As shown in Fig. 2, in the method using liner predictive coefficient parameters, four parameters, namely the parameter for the starting frame of the static segment in the syllable file for the preceding syllable 15, the parameter for the terminating frame of the static segment in the same syllable file 16, the parameter for the starting frame in the static segment in the syllable file for the independently uttered vowel same as the vowel of the preceding syllable 18, and the parameter for the terminating frame of the static segment in the same syllable file 19, are used for linear interpolation concatenating 15 to 16, 16 to 18, and 18 to 19.

In this way, according to the embodiment of this invention, the concatenation method using amp parameters has the advantage of there being little sense of discontinuity because the parameters for the preceding syllable are used unchanged until the terminating frame of the syllable for the prolonged sound section. In addition, the concatenation method using linear predictive coefficient parameters has the effect of being able to preserve phonological characteristics because the parameters from the syllable file for the same independently uttered vowel as the vowel in the preceding syllable are used in the prolonged sound section.

Next, another embodiment of this invention is described with reference to Fig. 3. Fig. 3 shows a concatenation method using amp parameters in another embodiment of this invention where a

THIS PAGE BLANK (USPTO)

prolonged sound is at the end of a word. In the figure, reference numeral 21 denotes the duration of a preceding syllable, 22 the duration of a prolonged sound when it is in the middle of a word (the same as the duration of a prolonged sound shown by 2, 12, and 32 in Figs. 1, 2, and 4), 23 a vowel segment of the preceding syllable, 24 a parameter for the starting frame of a static segment in the syllable file for the preceding syllable, 25 a parameter for the terminating frame of the static segment in the syllable file for the preceding syllable, 26 a prolonged sound section, 27 a parameter for the starting frame of a static segment in the syllable file for the prolonged sound section, and 28 a parameter for the terminating frame of the static segment in the syllable file for the prolonged sound section. In contrast to the case of Fig. 1, where the parameter would be kept at a constant value until the terminating frame of the static segment of the prolonged sound section 28, the amp parameter is kept at a constant value until the starting frame of the static segment of the prolonged sound section 27, and then subjected to linear interpolation so as to be set at 0 at the terminating frame of the static segment 28, because a prolonged sound is vocalized shorter at the end of a word than in the middle of the word. Note that, if linear predictive coefficient parameters are used in the same case, the concatenation method used is the same as when the prolonged sound is in the middle of a word (the concatenation method shown in Fig. 2).

By using the above word-end processing, it is possible to obtain perceptually natural prolonged sounds.

[Effect of the Invention]

According to this invention, it is possible to synthesize prolonged sounds that are perceptually extremely natural by a concatenation method using amp parameters and linear predictive coefficient parameters within ordinary syllable files used in CV (consonant-vowel) syllable concatenation, in rule-based synthesis methods, without creating special files for prolonged sounds.

THIS PAGE BLANK (USPTO)

4. Brief Description of Drawings

Fig. 1 shows concatenation method using amp parameters in an embodiment of this invention.

Fig. 2 shows a concatenation method using linear predictive coefficient parameters in an embodiment of this invention.

Fig. 3 shows a concatenation method using amp parameters in another embodiment of this invention where a prolonged sound is at the end of a word.

Fig. 4 shows a conventional concatenation method using amp parameters.

1, 11, 21: Duration of preceding syllable

2, 12, 22: Duration of prolonged sound

3, 13: Duration of subsequent syllable

4, 14, 23: Vowel segment of preceding syllable

5, 24: Amp parameter for starting frame of static segment in syllable file for preceding syllable

6, 25: Amp parameter for terminating frame of static segment in syllable file for preceding syllable

7, 17, 26: Prolonged sound section

8, 27: Amp parameter for starting frame of static segment in syllable file for prolonged sound section

9, 28: Amp parameter for terminating frame of static segment in syllable file for prolonged sound section

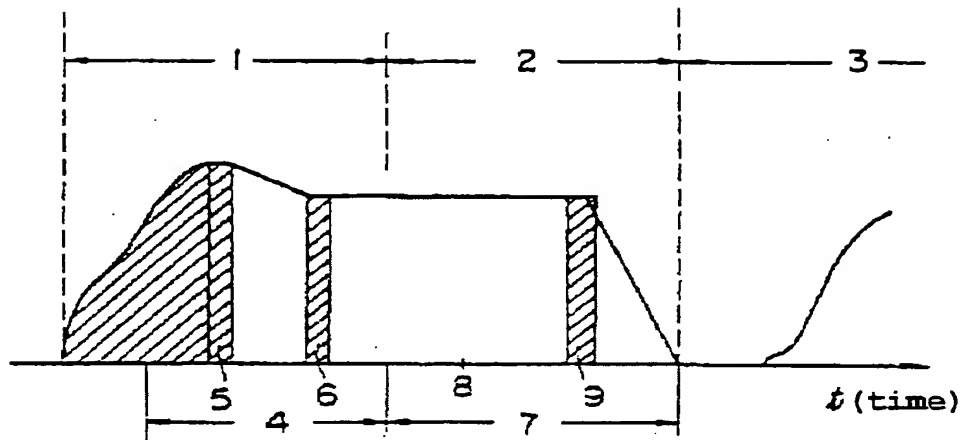
15: Linear predictive coefficient parameter for starting frame of static segment in syllable file for preceding syllable

16: Linear predictive coefficient parameter for terminating frame of static segment in syllable file for preceding syllable

18: Linear predictive coefficient parameter for starting frame of static segment in syllable file for independently uttered vowel

19: Linear predictive coefficient parameter for terminating frame of static segment in syllable file for independently uttered vowel

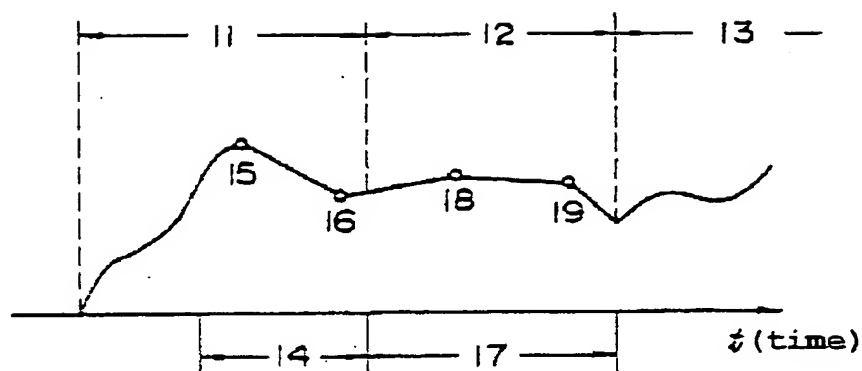
THIS PAGE BLANK (USPTO)



- 1: Duration of preceding syllable
- 2: Duration of prolonged sound
- 3: Duration of subsequent syllable
- 4: Vowel segment of preceding syllable
- 5: Amp parameter for starting frame of static segment in syllable file for preceding syllable
- 6: Amp parameter for terminating frame of static segment in syllable file for preceding syllable
- 7: Prolonged sound section
- 8: Amp parameter for starting frame of static segment in syllable file for prolonged sound section
- 9: Amp parameter for terminating frame of static segment in syllable file for prolonged sound section

FIG. 1

THIS PAGE BLANK (USPTO)



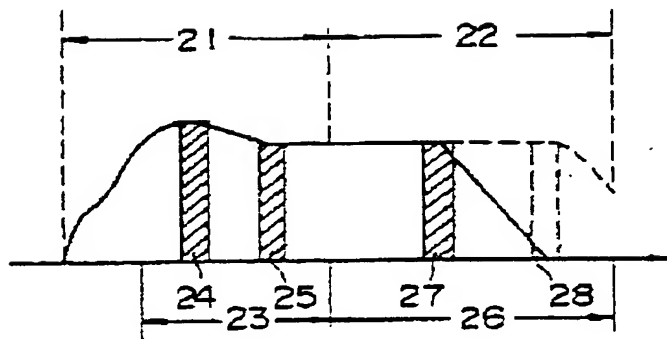
- 11: Duration of preceding syllable
- 12: Duration of prolonged sound
- 13: Duration of subsequent syllable
- 14: Vowel segment of preceding syllable
- 15: Linear predictive coefficient parameter for starting frame of static segment in syllable file for preceding syllable
- 16: Linear predictive coefficient parameter for terminating frame of static segment in syllable file for preceding syllable
- 17: Prolonged sound section
- 18: Linear predictive coefficient parameter for starting frame of static segment in syllable file for independently uttered vowel
- 19: Linear predictive coefficient parameter for terminating frame of static segment in syllable file for independently uttered vowel

FIG. 2

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

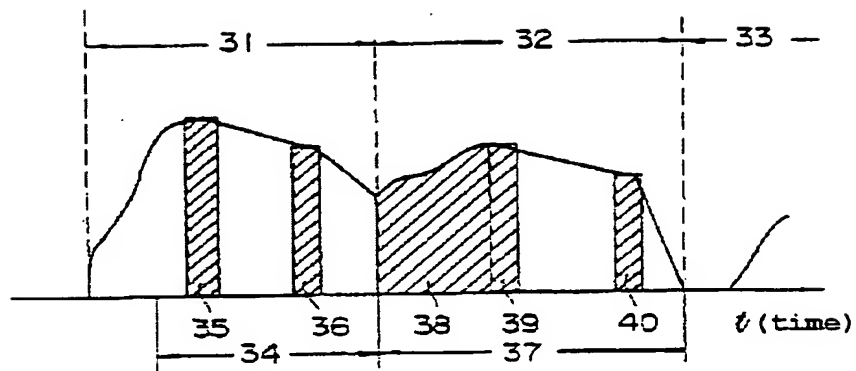
THIS PAGE BLANK (USPTO)



- 21: Duration of preceding syllable
- 22: Duration of prolonged sound
- 23: Vowel segment of preceding syllable
- 24: Amp parameter for starting frame of static segment in syllable file for preceding syllable
- 25: Amp parameter for terminating frame of static segment in syllable file for preceding syllable
- 26: Prolonged sound section
- 27: Amp parameter for starting frame of static segment in syllable file for prolonged sound section
- 28: Amp parameter for terminating frame of static segment in syllable file for prolonged sound section

FIG. 3

THIS PAGE BLANK (USPTO)



- 31: Duration of preceding syllable
- 32: Duration of prolonged sound
- 33: Duration of subsequent syllable
- 34: Vowel segment of preceding syllable
- 35: Amp parameter for starting frame of static segment in syllable file for preceding syllable
- 36: Amp parameter for terminating frame of static segment in syllable file for preceding syllable
- 37: Frame in which independently uttered vowel is embedded
- 38: Amp parameter from head of syllable file for independently uttered vowel to starting frame of static segment
- 39: Amp parameter for starting frame of static segment in syllable file for independently uttered vowel
- 40: Amp parameter for terminating frame of static segment in syllable file for independently uttered vowel

FIG. 4

THIS PAGE BLANK (USPTO)

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭63-25700

⑮ Int.Cl.⁴
G 10 L 9/14

識別記号 庁内整理番号
M-8622-5D

⑭ 公開 昭和63年(1988)2月3日

審査請求 未請求 発明の数 1 (全4頁)

⑬ 発明の名称 長音結合方法

⑯ 特 願 昭61-168134

⑰ 出 願 昭61(1986)7月18日

⑱ 発 明 者 小 林 理 恵 子 神奈川県横浜市港北区綱島東4丁目3番1号 松下通信工業株式会社内

⑲ 出 願 人 松下電器産業株式会社 大阪府門真市大字門真1006番地

⑳ 代 理 人 弁理士 星野 恒 司 外1名

明 細 書

1. 発明の名称 長音結合方法

2. 特許請求の範囲

後続音節に長音がかかる場合の長音結合方法において、アンブパラメータを使用する方法では、前音節のパラメータを長音区間音節の終端フレームまで一定値で結合し、線形予測係数パラメータを使用する方法では、長音区間に前音節の母音と同じ単独発声母音の音節ファイル中のパラメータを用い、これらのパラメータと前記前音節の母音の音節ファイル中の定常区間終端フレームのパラメータとを直線補間により結合することを特徴とする長音結合方法。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は、規則合成方式による音声合成方法における長音結合方法に関するものである。

(従来の技術)

従来の長音結合方法を第4図により説明する。

第4図はアンブパラメータを用いた従来の長音結合方法を示す図である。同図において、31は前音節の時間長、32は長音の時間長、33は後続音節の時間長、34は前音節中の母音部、35は前音節の音節ファイル中の定常区間始端フレームのパラメータ、36は前音節の音節ファイル中の定常区間終端フレームのパラメータを示す。37は前音節の母音と同じ単独発声母音を埋め込むフレームで、38(斜線部分)はこの単独発声母音の音節ファイル中の先頭から定常区間始端フレーム39までのパラメータを、39及び40は前記単独発声母音の音節ファイル中のそれぞれ定常区間始端および終端フレームのパラメータを示している。すなわち、従来の長音結合方法は、長音にあたる区間に前音節の母音と同じ単独発声母音の音節ファイルのパラメータを用い、他の音節の結合方法と同様に、これらのパラメータを前音節の音節ファイル中のパラメータとつなげている。

このように、上記従来の結合方法でも、長音記号にあたる区間に前音節の単独発声母音を置き換

特開昭63-25700 (2)

えれば、他の通常音節と同様のアルゴリズムで処理することができる。また、第4図はアンブパラメータを示しているが、線形予測係数パラメータも同様の方法で行っている。

(発明が解決しようとする問題点)

しかし、上記従来の方法では、長音にあたる区間に前音節の母音と同じ単独発声母音を使用して、聴感上長音らしく聞こえにくいという問題があった。

本発明は、規則合成方法によって長音を合成できる長音結合方法を提供するものである。

(問題点を解決するための手段)

本発明は、上記問題点を解決するために、アンブパラメータを用いる方法では、前音節のパラメータを長音区間音節の終端フレームまで一定値で結合し、線形予測係数パラメータを用いる方法では、長音区間に前音節の母音と同じ単独発声母音の音節ファイル中のパラメータを用い、これらのパラメータと前記前音節の母音の音節ファイル中の定常区間終端フレームのパラメータとを直線補

間により結合するものである。

(作用)

本発明は、上記方法により、音節ファイルに新たに長音ファイルを追加することなく、既存の音節ファイル中のパラメータを用いて、長音を含む合成音を合成するのに必要なパラメータ時系列を作成することができ、且つ聴感上自然に聞こえる長音が得られる。

(実施例)

本発明の一実施例を第1図及び第2図により説明する。第1図は本発明の一実施例におけるアンブパラメータによる結合方法を示す図、第2図は本発明の一実施例における線形予測係数パラメータによる結合方法を示す図である。第1図において、1は前音節の時間長、2は長音の時間長、3は後続音節の時間長、4は前音節の母音部、5は前音節の音節ファイル中の定常区間始端フレームのパラメータ、6は前音節の音節ファイル中の定常区間終端フレームのパラメータ、7は長音にあたる区間、8(時間軸上に点で示す)は長音区間の

音節ファイル中の定常区間始端フレームのパラメータ、9は長音区間の音節ファイル中の定常区間終端フレームのパラメータを示す。第1図に示すように、アンブパラメータによる方法では、前音節の終端フレーム6のパラメータをそのまま長音区間音節ファイル中の定常区間終端フレーム9まで一定値で用いる。

線形予測係数パラメータによる結合方法を示す第2図において、11は前音節の時間長、12は長音区間の時間長、13は後続音節の時間長、14は前音節の母音区間、15は前音節の音節ファイル中の定常区間始端フレームのパラメータ、16は前音節の音節ファイル中の定常区間終端フレームのパラメータ、17は長音区間、18は前音節の母音と同じ単独発声母音の音節ファイル中の定常区間始端フレームのパラメータ、19は単独発声母音の音節ファイル中の定常区間終端フレームのパラメータを示す。

第2図に示すように、線形予測係数パラメータによる方法では、前音節の音節ファイル中の定常

区間始端フレームのパラメータ15、同音節ファイル中の定常区間終端フレームのパラメータ16、前音節の母音と同じ単独発声母音の音節ファイル中の定常区間始端フレームのパラメータ18及び同音節ファイル中の定常区間終端フレームのパラメータ19の4点のパラメータを用いて、15から16、16から18、18から19を、各々直線補間で結合する。

このように、本発明の一実施例によれば、アンブパラメータによる結合方法では、前音節のパラメータをそのまま長音区間音節の終端フレームまで用いるので、聴感上の不連続感が少ないという利点を有する。また、線形予測係数パラメータによる結合方法では、長音区間に前音節の母音と同じ単独発声母音の音節ファイル中のパラメータを用いるので音節の特徴をそのまま保持できるという効果を有する。

次に、本発明の他の実施例を第3図により説明する。第3図は本発明の他の実施例における綴尾に長音がくる場合のアンブパラメータを用いた結合方法を示す図である。同図において、21は前音

特開昭63-25700 (3)

語の時間長、22は長音が語中にくる場合の時間長(第1,2,4図の2,12,32に示す長音の時間長と同じ)、23は前音節の母音区間、24は前音節の音節ファイル中の定常区間始端フレームのパラメータ、25は前音節の音節ファイル中の定常区間終端フレームのパラメータ、26は長音区間、27は長音区間の音節ファイル中の定常区間始端フレームのパラメータ、28は長音区間の音節ファイル中の定常区間終端フレームのパラメータを示す。第1図に示すと同様、アンプパラメータは、長音区間の定常区間終端フレーム28まで一定値を保持するところを、語尾に長音がある場合には、語中に長音がある場合より短か目に発声されるので、長音区間の定常区間始端フレーム27まで一定値を保持し、定常区間終端フレーム28のところで0と直線補間している。なお、同様なケースについて線形予測係数パラメータを用いる場合には、語中に長音がある場合と同様な結合方法(第2図に示す結合方法)を用いている。

上記のような語尾処理により、聴感上自然な長

音が得られる。

(発明の効果)

本発明によれば、特別に長音用ファイルを作成することなく、規則合成方式のうち、CV(子音、母音)音節結合で用いられる通常の音節ファイル中のアンプパラメータ及び線形予測係数パラメータを用いる結合方法により、聴感上極めて自然な長音を合成することができる。

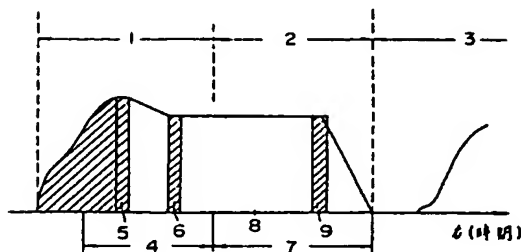
4. 図面の簡単な説明

第1図は本発明の一実施例におけるアンプパラメータによる結合方法を示す図、第2図は本発明の一実施例における線形予測係数パラメータによる結合方法を示す図、第3図は本発明の他の実施例における語尾に長音がある場合のアンプパラメータを用いた結合方法を示す図、第4図はアンプパラメータを用いた従来の長音結合方法を示す図である。

1, 11, 21 … 前音節の時間長、 2, 12, 22 … 長音の時間長、 3, 13 … 後続音節の時間長、 4, 14, 23 … 前音節の母音

区間、 5, 24 … 前音節の音節ファイル中の定常区間始端フレームのアンプパラメータ、 6, 25 … 前音節の音節ファイル中の定常区間終端フレームのアンプパラメータ、 7, 17, 26 … 長音区間、 8, 27 … 長音区間の音節ファイル中の定常区間始端フレームのアンプパラメータ、 9, 28 … 長音区間の音節ファイル中の定常区間終端フレームのアンプパラメータ、 15 … 前音節の音節ファイル中の定常区間始端フレームの線形予測係数パラメータ、 16 … 前音節の音節ファイル中の定常区間終端フレームの線形予測係数パラメータ、 18 … 単独発声母音の音節ファイル中の定常区間始端フレームの線形予測係数パラメータ、 19 … 単独発声母音の音節ファイル中の定常区間終端フレームの線形予測係数パラメータ。

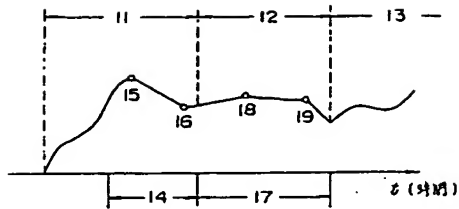
第1図



- 1… 前音節の時間長
- 2… 長音の時間長
- 3… 後続音節の時間長
- 4… 前音節の母音区間
- 5… 前音節の音節ファイル中の定常区間始端フレームのアンプパラメータ
- 6… 前音節の音節ファイル中の定常区間終端フレームのアンプパラメータ
- 7… 長音区間
- 8… 長音区間の音節ファイル中の定常区間始端フレームのアンプパラメータ
- 9… 長音区間の音節ファイル中の定常区間終端フレームのアンプパラメータ

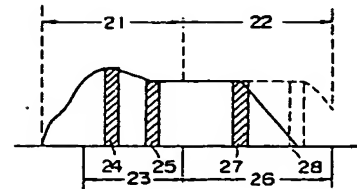
特開昭63-25700 (4)

第 2 図



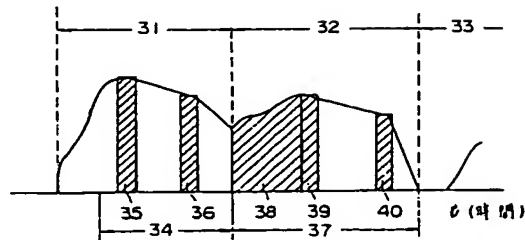
- 11… 前音節、時間長
- 12… 長音、時間長
- 13… 後続音節、時間長
- 14… 前音節、母音区間
- 15… 前音節、音節ファイル中、定常区間
始端フレーム、振幅予測係数バグメ-7
- 16… 前音節、音節ファイル中、定常区間
終端フレーム、振幅予測係数バグメ-7
- 17… 長音区間
- 18… 単独発声母音、音節ファイル中、定常区間
始端フレーム、振幅予測係数バグメ-7
- 19… 単独発声母音、音節ファイル中、定常区間
終端フレーム、振幅予測係数バグメ-7

第 3 図



- 21… 前音節、時間長
- 22… 長音、時間長
- 23… 前音節、母音区間
- 24… 前音節、音節ファイル中、定常区間
始端フレーム、アンパバグメ-7
- 25… 前音節、音節ファイル中、定常区間
終端フレーム、アンパバグメ-7
- 26… 長音区間
- 27… 長音区間、音節ファイル中、定常区間
始端フレーム、アンパバグメ-7
- 28… 長音区間、音節ファイル中、定常区間
終端フレーム、アンパバグメ-7

第 4 図



- 31… 前音節、時間長
- 32… 長音、時間長
- 33… 後続音節、時間長
- 34… 前音節、母音区間
- 35… 前音節、音節ファイル中、定常区間
始端フレーム、アンパバグメ-7
- 36… 前音節、音節ファイル中、定常区間
終端フレーム、アンパバグメ-7
- 37… 単独発声母音、埋め込みフレーム
- 38… 単独発声母音、音節ファイル中、定常区間
定常区間始端フレーム、アンパバグメ-7
- 39… 単独発声母音、音節ファイル中、定常区間
始端フレーム、アンパバグメ-7
- 40… 単独発声母音、音節ファイル中、定常区間
終端フレーム、アンパバグメ-7